

# BILINGUAL TEST AS A TEST ACCOMMODATION TO DETERMINE THE MATHEMATICS ACHIEVEMENT OF MAINSTREAM STUDENTS WITH LIMITED ENGLISH PROFICIENCY

**S. Kanageswari Suppiah Shanmugam<sup>1</sup>**

*Seameo Regional Center for Education in Science and Mathematics*

*Seameo Recsam, Penang*

&

**Ong Saw Lan**

*School of Educational Studies*

*Universiti Sains Malaysia*

<sup>1</sup>*Corresponding author: kanageswari@recsam.edu.my*

## ABSTRACT

**Purpose** – This study aims to investigate the validity of using bilingual test to measure the mathematics achievement of students who have limited English proficiency (LEP). The bilingual test and the English-only test consist of 20 computation and 20 word problem multiple-choice questions (from TIMSS 2003 and 2007 released items. The bilingual test consists of items in Malay and English languages. The English teachers classified their students into LEP and non-LEP groups.

**Methodology** – A total of 2,021 LEP and 2,747 non-LEP students from 34 schools were identified. Spiral administration was employed. A total of 2,399 students sat for the bilingual test and 2,369 for the English-only test. The scores were linked using RAGE-RGEQUATE version 3.22.

**Findings** – Findings revealed that in the word problem testlet, the bilingual test was one unit easier for LEP students, probably indicating that the Malay language adaptation assisted them. For the non-LEP students, the bilingual version did not provide any added advantage to them as they did not show elevated achievement score.

**Significance** – This study indicates bilingual test can be an appropriate test accommodation that validly measures both LEP and non-LEP students' mathematics achievement in countries where English is not the *native language but is used as the instructional language*.

**Keywords:** Computation testlet, word problem testlet, LEP, bilingual test, English-only test, mathematics achievement.

## INTRODUCTION

Successful mathematical learning encompasses the mastery of three broad domains of content, literacy and language skills (Robertson & Summerlin, 2005) which requires students to read and comprehend the language of the mathematical word problem and rewrite it to the abstract language of Mathematics denoted by relations and symbols (Irujo, 2007; Khisty, 1993). When two constructs like language and mathematical ability are so closely related, it is vital to ensure that the Mathematics test scores obtained by any student, mainly reflects his mathematical ability and that, the amount of that composite score due to his language ability is minimised (Abedi, Lord & Plummer, 1997). For limited English proficiency (LEP) students, language becomes a stumbling block because as they are learning the content knowledge, they are also learning the language used in the content (Lachat & Spruce, 1998; Virginia Department of Education, 2004). As such, for LEP students who are sitting for a Mathematics test, when testing is in English, the test scores may not be an accurate measurement of their mathematical ability.

One approach to gather true mathematics score ability is by providing test accommodation. Test accommodation refers to the changes adopted to either the test format or testing situation that do not alter the test construct but facilitates the test takers' language or physical deficiency by correcting the unfairness present within a disadvantaged group in order to address their unique needs without providing advantages to the general education student population (Wilde, 2007). Bilingual test is one type of test accommodation that provides LEP students with equal access to the test content as it removes the construct irrelevant variance and allows the original test items to be translated into the students' more proficient language and hence, reduces language barrier. Thus, putting the LEP and the non-LEP students on equal ground (American Educational Research

Association, American Psychological Association, & National Council on Measurement in Education AERA, 1999).

In the Malaysian context, this research explores the extent to which the bilingual test is an effective test accommodation that validly measure students' mathematics achievement, particularly in alleviating LEP students' linguistics complication. This is because there is the possibility that the bilingual test may benefit some students, have adverse effects on them or benefit only a particular group but detrimental to another.

## RESEARCH BACKGROUND

In Malaysia, majority of students whose first language is not English have limited English proficiency. English is a learnt language and teaching students in a language that is foreign to them aggravates the difficulty they face in understanding these subjects (Ain Nadzimah & Chan, 2003). Since English is Malaysian students' second or third language, they lack the level of English proficiency required to fairly demonstrate their mathematical skills especially in word problems (Fatimah Hashim & Zarina Ramli, 2004). Students with restricted English language background are subjected to enormous language impediments that threaten their ability to learn and perform effectively especially when English is the language of instruction (Wang, Reynolds & Walberg, 1991). In addition to LEP, lack of academic English also contributes to students' poor performance in Mathematics especially in word problem items (Brown, 2005).

## LITERATURE REVIEW

### *LEP Students' Mathematics Achievement*

Mathematics is particularly challenging to LEP students as it subsumes linguistic knowledge, conceptual knowledge and procedural knowledge. Linguistic knowledge is related to English proficiency while conceptual knowledge is based on the understanding of mathematical concepts that will direct students to select the correct operation. Procedural knowledge involves the different approaches of learning Mathematics that are defined by different cultures (Virginia Department of Education, 2004).

LEP students tend to exhibit significantly different performance in linguistically simplified mathematics items and linguistically non-simplified mathematics items, which suggests that language skills also contribute to their performance in addition to their mathematical skills (Abedi & Lord, 2001). According to Anderson (2007), the probability of obtaining a correct response is approximately the same for both groups of LEP and non-LEP for a non-linguistic mathematics item, unlike for a linguistically complex mathematics item where the LEP students are at a disadvantage compared to the non-LEP students. This is mainly because a successful mathematical solution is largely influenced by the language. This is because successful word problem solutions integrate a combination of multiple mathematical skills and linguistic and cognitive skills that occurs when students read and comprehend the text which is in a language foreign to them (Mather & Chiodo, 1994). Poor grasp of both the social language and academic mathematical language is a cause for poor students' performance in Mathematics (Ferari, 2004). The influence of social and academic languages can best be understood by exploring the interaction between 'Basic Interpersonal Communication Skills' (BICS) and 'Cognitive Academic Language Proficiency' (CALP) (Cummins, 2001). BICS is the students' familiar home language that they carry to school and CALP is the academic language of the subjects in school. The impact of the home language (BICS) and the academic language (CALP) can be studied using the following mathematics question.

"You have 20 dollars. You have 6 dollars more than me. How many dollars do I have?" (Baker, 1996, p. 151).

Students who attempt to solve the question using the academic language are able to obtain the correct answer as 14 dollars by conceptualising the question as 6 subtracted from 20. In contrast, students who interpret the question within the context of the home language will most likely get 26 as the wrong answer. This arises when students get confused with the meaning of the word 'more' used in a social setting usage which carries the connotation 'add up'. Comparatively, the home language can be developed within a shorter time period of two years of language learning, unlike the academic language that requires a longer duration of five to seven years and is dependent on factors like the intensity of language and the students' age (Collier, 1987).

### **Word Problem Items**

Unlike computation items, word problem items are ‘real world like’ problems placed within a context (Li, 1998; Nesher & Katriel, 1977). They are relatively denser with language and require multiple steps in solving when compared to computation items.

According to Oviedo (2005), students who attempt word problem items have to face the conflict that exists as a result of understanding the text embedded in the item, the context in which the item is placed and the problem solving strategies that are required for solving it. The mental struggle endured probably explains why students who can competently solve arithmetic computation do not necessarily display the same level of competency when solving word problem items (Oviedo, 2005; Valentin & Lim, 2004). Language proficiency appears to be a pre-requisite as it is necessary for students to unpack the text of the word problem items before solving the implied mathematical concepts using the correct mathematical operations.

In a study conducted in Malaysia by Lim (2002), Malay pupils from Chinese medium schools were found to be not doing well in mathematical word problems. The reason was they were weak in the Chinese language which was the academic language used in school but not their home language as their mother tongue was Bahasa Malaysia. The findings of this study imply that in order to solve word problems, students need to master a language well enough to understand the underlying mathematical concepts and skills. As reiterated by Reed (1999), students need to understand the text in the word problem items by removing the extraneous information posed by the linguistics features before selecting the correct arithmetic operation. The study also highlights the importance of mastering the language of instruction for academic excellence when compared to home language.

In another study, Bernardo (1999) reported that his findings indicated that Filipino-English bilingual students exhibited better performance in solving word problems that were written in their first language which was the Philippines language when compared to their performance in word problems written in English. Error analyses revealed that students’ ability to understand the text in the word problem was the reason behind their performance. Since

the word problems were written in their first language, they were able to display a higher degree of comprehension. In a later study which also involved Filipino-English bilinguals, he discovered that students with English as the first language outperformed their Philippines speaking counterpart. The likely reason was traced back to the education policy that mandates the compulsory usage of English as the language of instruction in Philippines. Therefore, the learning environment set in the English language benefitted students with English as their first language. The Filipino students with the Philippines language as the first language were not able to fully benefit from a classroom setting where lessons, materials, discussions and responses were conducted in the English language (Bernardo, 2005). The findings highlight that students tend to exhibit improved mathematical performance when teaching, learning and assessing are conducted in the language that they demonstrate higher proficiency.

As suggested by the findings of a study conducted by Clements (1980), limited proficiency poses serious complication in solving word problem items. He reported that approximately 50% of student's difficulty in solving word problem was associated with language. Among LEP students whose first language is not English, the linguistic difficulties are even more pronounced and normally, they resort to using all the resources in both the language of instruction and native language to help alleviate their linguistics shortcoming.

These studies reinforce language proficiency as a significant factor that contributes towards mathematics achievement especially for word problem items. They seem to suggest that proficiency in the language of instruction or the academic language used in school is more relevant than home language or native language proficiency.

### **Word Problem Model**

The Problem Model proposed by Kintsch and Greeno (Kintsch & Greeno, 1985) consists of the set of knowledge that is fed by the textual information of the word problem and the abstract problem model. To solve a mathematical problem, the mathematical knowledge is used together with a set of strategies that form a mathematical representation. This representation has two facets which are the text that provides the textual input and the problem

model that contains the necessary information fed from the text. The model distinguishes three different types of knowledge. The first translates the mathematical text into propositions while the second consists of a schema of mathematical relations that is used to build the problem model. The third involves the mathematical operations and computation skills that are essential in solving.

Therefore based on this problem model, to solve a mathematical problem a student infers the mathematical information embedded in the text by using their knowledge of the mathematical language and the language of the text, builds a conceptual representation and then selects the correct mathematical operations. As such, a successful mathematical solution is dependent on the students' academic language proficiency of the text, social language proficiency, mathematical language proficiency and mathematical competency.

### ***Bilingual Testing***

According to The National Association of the Education of Young Children (NAEYC) (2005), bilingual test booklet can provide an accurate measure of students' knowledge and skills. When students' proficiency in home language and English language could not be assessed, NAEYC (2005) proposed that students be assessed in both languages. These two languages are recommended to be used especially for mathematical word problem items (Jean, 2006). There are two options in administering a bilingual test. One is to produce two test booklets in each language and the teachers or the examinees will decide which booklet will be used for answering. Another option is to administer one booklet in two languages with one language version following another language version (Stansfield, 1997).

### ***Test Linking***

When administering two tests, the original and the translated tests must be linked so that the examinees' performance in the two forms can be compared (Rapp & Allalouf, 2003) as the differences in the scores could be due to the differences in the difficulties of the two test booklets, apart from the differences due to students' ability (Kolen & Brennan, 1995).

According to Holland and Dorans (2006), in order to link the scores of one test to the scores of another test, the scores of one test must

be transformed into the scores of the new form by using raw-to-scale conversion of the score from the old form. One of the designs that can be employed is the random equivalent-groups design which by using a spiralling process; the two tests are administered in an alternating manner. The advantages of this administration are that it reduces strain among students who otherwise have to sit for both forms of tests and as such, saves time. Another added advantage of random assignment of booklets is that it allows the teacher, class and school effect to be controlled ((Abedi, Courtney, Mirocha, Leon & Goldberg, 2005). The difference exhibited in the scores between the two groups is the result of the differences in the difficulty of the two test forms (Kolen & Brennan, 1995).

## RESEARCH OBJECTIVES

The focus of this research is to explore LEP students' mathematics performance in word problem items in bilingual test. However, to determine the effect of this accommodation, it is necessary to compare their performance in word problem testlet with the English-only test. In addition, their performance in the computation testlet, which eliminated language requirements, is also examined. In order to determine the effect of language ability on LEP student's mathematics achievement, the LEP students' performance is then compared with the non-LEP students' mathematics performance in the word problem testlet for both tests. Particularly, the objectives of this research are:

- a) To compare LEP students' mathematics performance in computation testlet in English-only and bilingual tests.
- b) To compare LEP students' mathematics performance in word problem testlet in English-only and bilingual tests.
- c) To compare non-LEP students' mathematics performance in word problem testlet in English-only and bilingual tests.

### ***Research Questions***

The research will answer several questions which are:

- a) Is there any difference in LEP students' mathematics performance in the computation testlet in the English-only and bilingual tests?

- b) Is there any difference in LEP students' mathematics performance in the word problem testlet for English-only and bilingual tests?
- c) Is there any difference in non-LEP students' mathematics performance in the word problem testlet for English-only and bilingual tests?

## RESEARCH SIGNIFICANCE

This research is of great significance in multiethnic, multicultural and multilingual societies like Malaysia. The Malaysian education system practices multilingualism at the primary level where there are three mediums of instructions which are Bahasa Malaysia (Malay language), Chinese language and Tamil language. As such, this research explores the validity of using bilingual test as a linguistics test accommodation to measure the mathematics achievement of both the LEP and non-LEP students who are the mainstream students. In countries where much research on test accommodation for LEP students have been conducted, the LEP students are not the mainstream students as they mainly consist of immigrants residing among English-native speakers. However, the scenario is different for this research. The bilingual test in this research will be administered in the instructional language (English) and the national language (Bahasa Malaysia) to determine the mathematical ability of the mainstream students who particularly face compounding language difficulties in the instructional language. Hence, the findings will shed some light on bilingual test as a valid test accommodation that can measure particularly, LEP students' mathematics achievement by removing language which has become the construct irrelevant variance.

The findings of this study will provide great insight on the appropriateness and utility of bilingual test to accurately measure LEP students' mathematics achievement. This is because this type of accommodation may benefit the general students or have adverse effects on them or benefit one group but detrimental to another.

In addition, the testlets in the bilingual test may even favour or disfavour certain groups of students as word problem testlets have more language load when compared to the computation testlet. Hence, the findings will provide an avenue to identify groups of

students which may suffer or benefit from the different testlets. On a broader scale, findings from this study can provide useful information to countries which are similar to Malaysia that does not have English as their native language in designing their assessment policy.

## **LIMITATIONS AND DELIMITATIONS**

The main limitation is in the selection of schools as the sample of this research. The schools selected must meet the condition where the mathematics teachers had completed all the topics that were included in the test content before the test administration during the first week of October. This limitation was imposed because administering the test after October was impossible as it would coincide with the school year end examination and followed by the year end school holidays before the new school year commenced.

In addition, the mathematics test items that were used consisted of only multiple choice questions (MCQ). Other types of test items like the constructed-response format were not used as the responses would depend on the students' higher order of reading, thinking and writing skills (Schulte, Elliot & Kratochwill, 2001) due to the intense linguistic density and as such, could form another new study of its very own.

Another limitation is that even though the researcher took extra steps of holding briefings, and a one-to-one session with the English teachers about the LEP designation, there was no objective way of checking whether there was any misclassification due to error in LEP designation.

## **METHODOLOGY**

### ***Research Design***

The research design employed is the random equivalent group design with spiral administration. For each school, six classes were selected. In a class, each student was assigned with either the English-only or the bilingual tests. By alternating each test booklet, all the

students in the class sat for one Mathematics test. LEP students were identified by using the expert judgment of the English teachers who taught at the respective classes.

### ***Mathematics Test***

This study utilised grade eight released Mathematics items from TIMSS 1999 and 2003 (International Association for the Evaluation of Educational Achievement, 2007). Three committees of teachers were set up to help construct the test. The first committee that was assembled consisted of three Form One and Form Two bilingual Malaysian Mathematics teachers with the experience of at least five years in teaching Form One and Form Two Mathematics. They selected the items that denoted the learning outcomes defined in the Curriculum Specifications Mathematics Form One (Curriculum Development Centre, 2003a) and Curriculum Specifications Mathematics Form Two (Curriculum Development Centre, 2003b) which were published by the Malaysian Curriculum Development Centre (CDC). They also classified them as computation or word problem items whereby word problem item was defined as ‘real world like’ problems placed within a context. The second committee is a committee of three teacher-translators with the experience of at least five years, who possessed language proficiency in both English and Malay languages and have assessment literacy. Their role was to adapt the English test items into the Malay language. The final committee consisted of another independent verification committee of 20 Form One and Form Two bilingual Malaysian Mathematics teachers with the experience of a minimum of five years in teaching Form One and Form Two Mathematics. They examined the test items for content validity.

There were 105 MCQ items chosen from the TIMSS released items, with 49 items from TIMSS 1999 and 94 from TIMSS 2003. The first teacher committee deduced that these items spread across all the forms from Form One to Form Five. Among these items, 62 of them were identified as representing the content domain included in the Form One and Form Two syllabi. The item selection criteria was primarily driven by content validity and the suitability of items for Form Two students. From the 62 items, 30 were word problem items and 32 were computation items. The items were pilot tested. The data from pilot testing was used to compute item difficulty,

item discrimination and the Cronbach's Alpha internal consistency reliability coefficient to provide further statistical evidence in item selection. The final mathematics test consists of 40 items (20 computation and another 20 word problem items) were used for the actual study. However, the decision on inclusion of items was on well-representativeness of the test content (Tinkelman, 1971), that is, content validity overruled statistical properties which merely served as a guide to detect poor items (Henrysson, 1971).

These 40 items were arranged by topics with computation items appearing first before the word problem items to form the layout for the two tests. The English-only version of the Mathematics test had two sections. The first section included the personal particulars of the subjects, mainly their race and gender. The second section consisted of 40 MCQ that formed the Mathematics test. These items have either three or four distracters with only one correct response. The Malay and English bilingual test consisted of three sections. The first section addressed students' personal particulars while the second section consisted of the 40 mathematics items. Similar to the English-only version, the second section had the same items, number of items and layout. However, an additional feature was the Malay language version. For each item and the accompanying diagrams, the Malay language version appeared immediately after the English version, in a square parenthesis using bold italic print of the same font size. The third section sought information about the utility of the bilingual version.

### ***Sample***

The student sample design that was employed was a two-stage cluster sampling of schools at the first stage and classroom at the second stage. Cluster sampling was adopted at each stage. From the 108 secondary schools in Penang, 29 schools were selected with 17 schools located in the Penang Island and another 12 schools from Penang mainland while only five schools in the Perak state were selected due to distance and mainly time constraints. Schools were selected by using purposive sampling because only schools that had completed the Form Two Mathematics syllabus by the first week of October could be used as the test content covered all these mathematics topics.

Six classes were chosen to represent the high, intermediate and low abilities students for an even student distribution. However, in schools with less than six classes, all the classes were selected. In this study, 12 schools had less than six classes. A total of 4,768 students sat for the tests.

LEP students were identified using teacher's judgment as teachers teaching the English subject were in the best position to provide expert judgment on classifying LEP students (Cummins, 1984). The English teachers of the classes concerned were requested to use their expert judgment to classify the students as LEP or non-LEP by using the class name list. These English teachers' expertise were used because they knew their students well and were in the best position to offer unbiased professional judgments about their students' English language proficiency. To ensure an objective LEP classification, a descriptive guideline of LEP and non-LEP was given to each teacher during a short briefing that was conducted between the English teachers and the researcher. This guideline was meant to assist them in the LEP designation as it provided a description related to the oral and written skills that defined LEP and non-LEP students. This guideline was prepared based on the descriptions given for the different bands used in scoring written English tests and oral assessments in national examinations. This guideline was validated by five English language teachers who agreed with the LEP and non-LEP designation. Based on this guideline, 2,021 LEP students and 2,747 non-LEP students were identified from 4,768 students with 2,399 students who sat for the bilingual test and 2,369 for the English-only test. The difference of 30 between the numbers of students who sat for the two tests was due to 30 English-only test booklets that were distributed to the students during the test administration were left unanswered.

Before the test booklets were distributed, the students were briefed for 15 minutes. The teacher-invigilator read aloud to the students a set of written instructions that was given to him. This was done to ensure that the students were aware of the test administration procedures. Students were reminded that the scores obtained from the test would not be used for any school related assessments. The languages used during the briefing were Malay and English. They were given time to fill in their personal particulars before the test started. They were

allowed one hour to answer the items. Students who had received the bilingual test booklet were particularly reminded to use the one hour to answer only the second section as they would be given an additional ten minutes to answer the third section.

## DATA ANALYSIS

Before analysing the data, the items were given dichotomous score which was either correct or incorrect. The correct response was given the score '1' while an incorrect response was given score '0'. Unanswered items were treated as incorrect and therefore was assigned score '0'. The scores of the 20 computation items were added to represent the students' mathematics achievement for the computation testlet while the scores of the 20 word problem items were added to represent the students' mathematics achievement for the word problem testlet.

To link the scores RAGE-RGEQUATE (Version 3.22) by Kolen (2005) was used. This software uses the postsmoothed equipercentile method for random group. For equipercentile equating, a cumulative distribution function of scores on the new form that is converted to the old form scale is plotted. By using this curve, scores on the new form which is the bilingual test was identified by determining the scores of the old form that shared the same percentile rank. This curve caters to the differences in the booklets difficulties and the difficulties may vary for different score range (Kolen & Brennan, 1995). The bilingual test is the new form while the English-only test is the old form.

In the first analysis, the LEP students' mathematics achievement in the computation testlet was compared between the English-only test and the bilingual test after the scores were linked for both tests. In the second analysis, the LEP students' mathematics achievement in the word problem testlet was compared between the English-only test and the bilingual test and the analysis is repeated for the non-LEP students in the final round of analysis. In addition, students' responses to the usefulness of the Malay language version in the bilingual test was analysed by documenting their written comments. SPSS version 16.00 was also used to compute the percentage of students who used the Malay translation for each item.

## FINDINGS

### ***LEP students' mathematics performance in computation testlets in English-only and bilingual tests***

To examine the LEP students' mathematics performance in the computation testlet in the English-only and bilingual tests, the first round of test linking analysis was conducted. Table 1 exhibits the output.

Table 1 shows that  $S=0.01$  represents the best approximation for the linked form as this value is the closest to the values of the mean, standard deviation, skewness and kurtosis for the English-only test. The difference between the mean, standard deviation, skewness and kurtosis are 0.003, 0.0028, 0.0051 and 0.0015 respectively. Using  $S=0.01$ , the linked scores for both tests are as displayed in Table 2.

Table 1

#### *Postsmoothing Raw Score Moments for LEP Students' Score in the Computation Testlet*

Test Form	Mean	Standard Deviation	Skewness	Kurtosis
English-only	10.1471	4.5222	0.1753	2.0531
Bilingual	10.1487	4.5170	0.1219	2.1073
Unsmoothed	10.1480	4.5177	0.1766	2.0497
$S=0.01:$	10.1441	4.5194*	0.1702	2.0516*
$S=0.05:$	10.1453	4.5151	0.1705*	2.0617
$S=0.10:$	10.1476*	4.5087	0.1687	2.0781
$S=0.20:$	10.1616	4.4965	0.1467	2.1079
$S=0.30:$	10.1782	4.4950	0.1216	2.1087
$S=0.40:$	10.1782	4.4950	0.1216	2.1087
$S=0.50:$	10.1782	4.4950	0.1216	2.1087
$S=0.75:$	10.1782	4.4950	0.1216	2.1087
$S=1.00:$	10.1782	4.4950	0.1216	2.1087
Linear:	10.1471	4.5222	0.1219	2.1073

As displayed in Table 2, there were no changes in the score for the computation testlet for the LEP students when the scores of the two tests were linked. This indicates that they exhibited similar mathematics performance in the computation testlet for both tests.

A probable reason could be due to the minimal language load present in the computation testlet that reduced their dependency on the translated version to decipher the mathematical content of the items.

Table 2

*Postsmoothing Raw-to-Raw Score Conversion for LEP Students' Score in Computation Testlet*

Bilingual Score	Unrounded English Score	Rounded English Score	Bilingual Score	Unrounded English Score	Rounded English Score
0	0.07	0	11	11.00	11
1	1.21	1	12	12.04	12
2	2.22	2	13	13.12	13
3	3.19	3	14	14.14	14
4	4.20	4	15	15.07	15
5	5.11	5	16	15.97	16
6	5.96	6	17	16.97	17
7	6.81	7	18	18.05	18
8	7.73	8	19	19.10	19
9	8.78	9	20	20.03	20
10	9.92	10			

*LEP students' mathematics performance in word problem testlet in English-only and bilingual tests*

To explore the LEP students' mathematics performance in the word problem testlet in the English-only and bilingual tests, another test linking analysis was conducted. Table 3 displays the results that were obtained.

Table 3

*Postsmoothing Raw Score Moments for LEP Students' Score in Word Problem Testlet*

Test Form	Mean	Standard Deviation	Skewness	Kurtosis
English-only	7.4434	4.0384	0.6928	2.8584
Bilingual	7.8170	4.2364	0.5625	2.5938
Unsmoothed	7.4407	4.0301	0.6929	2.8402

(continued)

Test Form	Mean	Standard Deviation	Skewness	Kurtosis
S=0.01:	7.4383*	4.0382*	0.6696*	2.8172*
S=0.05:	7.4378	4.0358	0.6673	2.8092
S=0.10:	7.4383	4.0378	0.6673	2.8003
S=0.20:	7.4542	4.0508	0.6449	2.7328
S=0.30:	7.4718	4.0542	0.6140	2.6801
S=0.40:	7.4834	4.0557	0.5933	2.6490
S=0.50:	7.4927	4.0569	0.5768	2.6253
S=0.75:	7.4991	4.0584	0.5667	2.6106
S=1.00:	7.4991	4.0584	0.5667	2.6106
Linear:	7.4434	4.0384	0.5625	2.5938

From Table 3, S=0.01 represents the best approximation to the values of the mean, standard deviation, skewness and kurtosis for the linked form for the word problem testlet as the difference noted between the mean, standard deviation, skewness and kurtosis are 0.0051, 0.0002, 0.0232 and 0.0412 respectively. Using S=0.01, the scores for both tests were linked as shown in Table 4 while the summary of the linked scores is as shown in Table 5.

Table 4

*Postsmoothing Raw-to-Raw Score Conversion for LEP Students' Score in Word Problem Testlet*

Bilingual Score	Unrounded English Score	Rounded English Score	Bilingual Score	Unrounded English Score	Rounded English Score
0	0.14	0	11	10.38	10
1	1.32	1	12	11.20	11
2	2.01	2	13	12.15	12
3	2.84	3	14	13.31	13
4	3.84	4	15	14.41	14
5	4.82	5	16	15.43	15
6	5.74	6	17	16.48	16
7	6.60	7	18	17.77	18
8	7.46	7	19	18.87	19
9	8.43	8	20	19.96	20
10	9.47	9			

Table 5

*Summary of Raw-to-Raw Score Conversion for LEP students' Score in the Word Problem Testlet*

Raw Score Range for Bilingual Test	LEP Students' Linked Score for English-only Test
0-7	No Changes
8-17	-1
18-20	No Changes

From Tables 4 and 5, it can be deduced that the LEP students' score for the word problem testlet in the English-only test dropped by one unit indicating that the LEP students exhibited better mathematics performance in bilingual test when compared to the English-only test. This probably was because they were tapping on both language resources, mainly the Malay language to unpack the language loads of the mathematics items. However, the students in the upper and lower end score range were not affected. This form of observation is anticipated as students at the lower end lack extensive mathematical mastery that language may no longer be a contributing factor for successful mathematical solution while the students at the upper end are cognitively capable and have mastered the mathematics and language components.

*Non-LEP students' mathematics performance in word problem testlets in English-only and bilingual tests*

To investigate the non-LEP students' mathematics performance in the word problem testlet in the English-only and bilingual tests, another test linking analysis was conducted. Table 6 shows the output.

From Table 6,  $S=0.01$  is the most suitable value that gives the smallest difference for mean, standard deviation, skewness and kurtosis between the two tests. The difference observed between the mean, standard deviation, skewness and kurtosis are 0.0004, 0.0015, 0.0016 and 0.0041 respectively. By using  $S=0.01$ , the scores were linked as shown in Table 7.

Table 6

*Postsmoothing Raw Score Moments for non-LEP Students' Score in the Word Problem Testlet*

Test Form	Mean	Standard Deviation	Skewness	Kurtosis
English-only	12.2412	4.4261	-0.3184	2.1392
Bilingual	12.2156	4.5196	-0.3280	2.1248
Unsmoothed	12.2427	4.4226	-0.3159	2.1311
S=0.01:	12.2408*	4.4276*	-0.3200*	2.1433*
S=0.05:	12.2471	4.4371	-0.3271	2.1316
S=0.10:	12.2495	4.4396	-0.3293	2.1294
S=0.20:	12.2495	4.4396	-0.3293	2.1294
S=0.30:	12.2495	4.4396	-0.3293	2.1294
S=0.40:	12.2495	4.4396	-0.3293	2.1294
S=0.50:	12.2495	4.4396	-0.3293	2.1294
S=0.75:	12.2495	4.4396	-0.3293	2.1294
S=1.00:	12.2495	4.4396	-0.3293	2.1294
Linear:	12.2412	4.4261	-0.3280	2.1248

From Table 7, it can be noted that there were no changes in the scores after linking the two tests. The non-LEP students' mathematics performance in the word problem testlet remained unchanged for both English-only and bilingual tests, most probably because they possessed sufficiently strong foundation in the English language that enabled them to understand the text embedded in the word problem items that were richly loaded with language. The results suggest that the bilingual version did not elevate the non-LEP students' mathematics performance by unnecessarily providing extra language support.

Table 7

*Postsmoothing Raw-to-Raw Score Conversion for non-LEP Students' Score in Word Problem Testlet*

Bilingual Score	Unrounded English Score	Rounded English Score	Bilingual Score	Unrounded English Score	Rounded English Score
0	0.02	0	11	11.02	11
1	1.06	1	12	12.01	12
2	2.11	2	13	12.99	13

(continued)

Bilingual Score	Unrounded English Score	Rounded English Score	Bilingual Score	Unrounded English Score	Rounded English Score
3	3.16	3	14	13.98	14
4	4.20	4	15	14.96	15
5	5.23	5	16	15.94	16
6	6.22	6	17	16.92	17
7	7.17	7	18	17.92	18
8	8.12	8	19	18.94	19
9	9.08	9	20	19.97	20
10	10.04	10			

***Students' responses on the utility of the Malay translation in the bilingual test***

In order to get a better description on the usefulness of the Malay adaptation, students were asked to respond to three questions in the third section of the bilingual test. They were to indicate whether they found the Malay translation useful or not. They were also asked whether they used the Malay translation to help them understand the items and if they responded positively, they would need to write out the question number of those items. Based on these responses, the percentage of students who relied on the Malay translation was calculated. Students' written comments were also documented.

From 2399 students who answered in the bilingual test, 43.7% were LEP students while 56.3% were non-LEP students. From the analyses, 89.5% of LEP students and 70.9 % of the non-LEP students found the Malay translation helpful. 88.6% of the LEP students and 67.1% of the non-LEP students used the Malay version to understand the questions. Among the LEP students, 92% used the Malay translation to answer all the items while 74% of the non-LEP students used the Malay language translation to ease understanding. From these figures, it can be deduced that LEP students relied more on the Malay translation as a source of understanding for both the computation and word problem items when compared to their counterpart. The percentage of LEP students who relied on the Malay translation for the items ranged from 32.9 to 53.3 while for the non-LEP students, the range was within 23.9% to 51.8%. However, there seem to be no significant pattern between the percentage of students who used the Malay translation for the computation and word problem items. An interesting finding from this study is that some of the LEP

students claimed that they did not find the Malay translation useful. Their simple comments like “teacher teach in English, so I answer in English” shows that they were able to comprehend the mathematical concepts of the items presented in the English language as the language of instruction during their Mathematics lesson was also in English. Since the language of assessment matched the language of instruction, there was no need to rely on the Malay language translation even though to some of them, it was their mother tongue. Another observation is that LEP students who received their primary education from Chinese schools did not find the bilingual test especially the Malay translation helpful. They believed that the Chinese language translation will be more helpful and useful to them when compared to the Malay language translation. This is because the language of instruction in their primary school was the Chinese language and later on was switched to the English language in the secondary school. Therefore, the Malay translation did not offer much assistance as it was not the language of instruction at both levels of their education.

## CONCLUSION AND DISCUSSION

Bahasa Malaysia is generally the dominant language of Malaysian students as it is the national language that unites all Malaysians. Thus, the bilingual version with the Malay translation provides the language support for LEP students who are in dire need of linguistics assistance. By relying on the Malay translation, the LEP students were more able to arrest the language load of the word problem items. This observation can only be extended to LEP students whose medium of instruction in the primary school was the Malay language. The findings also indicate that generally LEP students did not depend on the translated Malay language version for the computation testlet. Therefore, bilingual test as a test accommodation is capable of alleviating LEP students’ language impediments.

For the non-LEP students who are capable of unpacking the English language load in the mathematics items in both tests, the bilingual test did not provide them with unnecessary language advantage in the word problem testlet. Bilingual test as a test accommodation would have been erroneous if it had been advantageous to them because then it will be unfair to the LEP students whose scores will then be pitched at a different take off value.

Since the findings of this study suggest that the bilingual test helped the LEP students to tolerate with the language load but did not unnecessarily offer advantage to the non-LEP students, it can be concluded that bilingual test is a valid test accommodation. As recapitulated by Sireci (1997), valid test accommodations will not put any one group at an advantage or disadvantage as was the case of the bilingual test in this study.

The LEP students' comments which suggested their preference to answer in the language used during instruction which was English and the high percentage of LEP students who used the Malay version to understand the items seem to indicate that reducing the language load may help LEP students. These students most probably used the bilingual items in Malay to comprehend the items written in English as was confirmed by a study conducted by Duncan et al. (2005). The students in their study used the bilingual test items to overcome their language impairment. As such, simplifying the structure of the English language in a mathematics item may be a valid assessment procedure when designing a Mathematics test.

In addition, LEP students from Chinese primary schools claimed that the Malay translation in the bilingual test was not particularly helpful. Since the language of instruction in the primary schools was the Chinese language and English at the secondary level, these students did not find the Malay version useful. Even though they learned Malay language as a compulsory subject and are familiar with the language, it was not helpful in enhancing their understanding of the mathematical terminologies.

This phenomenon can be explained by using Cummins' (2001) CALP and BISC. These LEP students did not face difficulty understanding the mathematical terminologies which was learnt in Chinese or English as these languages were adopted as the academic language at the primary and secondary levels. However, Malay is not their mother tongue language or their home language. As it was also not the academic language, the Malay language version was not helpful to these students. It is important to be aware that the Malay language version did not help them to alleviate the linguistic impediments as it was not their academic language or their home language. It is certainly not because it was not one of their dominant languages as Bahasa Malaysia is the national language and the language of communication among Malaysians.

In countries where English is not the native language but is used as the medium of instruction, bilingual test helps LEP students to demonstrate their true mathematics achievement as it addresses their linguistics disability. As Bernardo (2005) stated, students whose first language is not English are at a disadvantage when participating in a classroom discourse that is enriched by resources and discussions where teachers and students engage in English. Only when they are able to understand the text that embodied the items can they project better mathematics performance (Bernardo, 1999).

While interpreting the results of this study, it is important to interpret these findings within the context of this study where the language of instruction was in English and the items that were used were also in the English language. This finding should not be generalised to future policy which may revert the language of instruction for Mathematics to the Malay language. This is because the Malay language is the national language and is understood by all Malaysian students as it is formally taught in all Malaysian schools.

## IMPLICATIONS AND RECOMMENDATIONS

The implication is that minimising the language load of mathematics items can be the answer to measure students' true mathematics achievement, especially in word problem items. One way of achieving this is by using diagrams for items that are rich in text. During test construction phase, teachers should be aware of the pressing need to construct linguistically simplified mathematics items that reduce or remove the unnecessary language load particularly in problem solving tasks. This care should also be exercised at every stage of test construction which also includes writing of test instructions.

## REFERENCES

Abedi, J., & Lord, C. (2001). The language factor in Mathematics test. *Applied Measurement in Education, 14*(3), 219-234.

Abedi, J., Courtney, M., Mirocha, J., Leon, S., & Goldberg, J. (2005). *Language accommodations for English language learners in large-scale assessments: Bilingual dictionaries and linguistic modification*. CSE Report 666. Retrieved from <http://www.cse/ucla.edu/products/reports/r666.pdf>

Abedi, J., Lord C., & Plummer, J. R. (1997). *Final report of language background as a variable in NAEP mathematics performance* (CSE Tech. Rep. No. 429). Los Angeles: University of California, Center for the Study of Evaluation/National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from <http://cunningham.acer.edu.au/inted/onlinedocs/1999/112897.pdf>

Ain Nadzimah Abdullah & Chan, S. W. (2003). Gaining linguistic capital through a bilingual language policy innovation. *South Asian Language Review. XIII*, 1-2.

American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (1999). *Standard for Educational and Psychological Testing*. Washington, DC: American Psychological Association

Anderson, J. (2007). *The Mathematics of language*. Harvard Graduate School of Education. Retrieved from [http://www.gse.harvard.edu/news\\_events/features/2007/03/22\\_martiniello.php](http://www.gse.harvard.edu/news_events/features/2007/03/22_martiniello.php)

Ansley, T. N., & Forsyth, R. A. (1990). An investigation of the nature of the interaction of reading and computational abilities in solving Mathematics word problems. *Applied Measurement in Education, 3*(4), 319-329.

Baker, C. (1996). *Foundations of bilingual education and bilingualism*. Great Britain: Bath Press.

Bernardo, A. B. I. (1999). Overcoming obstacles to understanding and solving word problems in mathematics. *Educational Psychology, 19*, 149-163.

Bernardo, A. B. I. (2005). Language and modelling word problems in Mathematics among bilinguals. *The Journal of Psychology, 139*(5), 413-425.

Brown, C.L. (2005). Equity of literacy-based Math performance assessment for English Language Learners. *Bilingual Research Journal, 29*(2), 337-364.

Clements, M. A. K. (1980). Analysing children's errors on written mathematical tasks. *Educational Studies in Mathematics, 11*, 1-21.

Collier, V. (1987). Age and rate of acquisition of second language for academic purposes. *TESOL Quarterly, 21*, 617-641.

Cummins, J. (1984). *Bilingualism and special education: Issues in assessment and pedagogy*. Great Britain: Short Run Press Ltd. from <http://www.iteachilearn.com/cummins/bicsalp.html>

Curriculum Development Centre. (2003a). *Curriculum specifications mathematics form one*. Ministry of Education, Dewan Bahasa dan Pustaka: Kuala Lumpur, Malaysia.

Curriculum Development Centre (2003b). *Curriculum specifications mathematics form two*. Ministry of Education, Dewan Bahasa dan Pustaka: Kuala Lumpur, Malaysia.

Duncan, T.G., et al., (2005). Study of a dual-language test booklet in eighth-grade Mathematics. *Journal of Applied Measurement in Education*, 18(2), 129-161.

Fatimah Hashim & Zarina Ramlan (2004). Language literacy impacting on mathematical literacy: Challenges for low proficiency ESL learners. *ETY Journal* 5(2), 11-14.

Ferari, P.F. (2004). *Mathematical language and advanced Mathematics learning*. Proceedings of the 28<sup>th</sup> Conference of the International Group for the Psychology of Mathematics Education, July 14-18, Bergen, Norway, 2, 383-390.

Gierl, M. J., & Khaliq, S.H. (2001). Identifying sources of differential item and bundle functioning on translated achievement tests: A confirmatory analysis. *Journal of Educational Measurement*, 38 (2), 164-187.

Hambleton, R. K., & Patsula, L. (1999). *Increasing the validity of adapted tests: Myths to be avoided and guidelines for improving test adaptation practices* [Electronic version]. Retrieved from <http://www.testpublishers.org/journal01.htm>

Henrysson, S. (1971). Gathering, analyzing, and using data on test items. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 130-159). Washington, D.C.: American Council on Education.

Holland, P. W., & Dorans, N.J. (2006). Linking and Equating. In R. L. Brennan (Ed.), *Educational measurement* (pp. 187-220). United States of America: Praeger Publishers.

International Association for the Evaluation of Educational Achievement. (2007). *TIMSS-USA- 8 Mathematics items*. Retrieved 11 November, 2007, from [http://nces.ed.gov/Timss/pdf/IIMSS\\_Mathh\\_Items.pdf](http://nces.ed.gov/Timss/pdf/IIMSS_Mathh_Items.pdf)

Irujo, S. (2007). Teaching Math to English Language Learners: Can research help? *ELL Outlook*. Mac/April from [http://www.coursescrafters.com/ELL-Outlook/2007/Mar\\_apr/ELL\\_OutlookITIN](http://www.coursescrafters.com/ELL-Outlook/2007/Mar_apr/ELL_OutlookITIN)

Jean, I. J. (2006). *Relationship between degree of bilingualism and mathematical word problem solving*. Retrieved from [http://www.ed.gov/rschstat/research/pubs/oieresearch/conference/jean\\_200601.pdf](http://www.ed.gov/rschstat/research/pubs/oieresearch/conference/jean_200601.pdf)

Khisty, L. (1993). *A naturalistic look at language factors in Mathematics teaching in bilingual classroom*. Paper presented at Third National Research Symposium on Limited English Proficient Student Issues: Focus on Middle and High School Issues, Washington, DC: U.D. Department of Education, Office of Bilingual and Minority Language Affairs. Retrieved from <http://www.ncela.gwu.edu/files/rcd/BE019310/A%20Naturalistic%20Look.pdf>

Kintsch, W., & Greeno, J. G. (1985). Understanding and solving word arithmetic problems. *Psychological Review*, 92(1), 109-129.

Kolen, M. J. (2005). *RAGE-RGEQUATE* (Version 3.22) [Computer software]. The University of Iowa, Iowa City: Iowa Testing Programs.

Kolen, M. J. & Brennan, R. L. (1995). *Test equating: Methods and practices*. Iowa City, USA: Springer-Verlag New York, Inc.

Lachat, M. A., & Spruce, M. (1998). *Assessment reform, equity, and English language learners* - Office of Educational Research and Improvement, Northeast and Island Regional Educational Laboratory at Brown University. Retrieved from <http://www.alliance.brown.edu/pubs/asellblb.pdf>

Li, Y. (1998). *The role and use of word problems in school Mathematics*. *Edumath*, 6. Retrieved from [http://www.fed.cuhk.edu.hk/~flee/mathfor/edumath/9806/10li\\_yp.html](http://www.fed.cuhk.edu.hk/~flee/mathfor/edumath/9806/10li_yp.html)

Lim, C. S. (2002). *Cultural differences and mathematics learning in Malaysia* [Electronic version]. Retrieved from <http://www.math.admu.edu/tsg22/lim.html>

Mather, R. C. & Chiodo, J.J. (1994). A mathematical problem: How do we teach Mathematics to LEP elementary students? *The Journal of Educational Issues of Language Minority Students*, 13, 1-12.

National Association for the Education of Young Children. (2005). *Screening and assessments of young English-language learners*. Retrieved from [http://www.naeyc.org/about/positions/ELL\\_Supplement.asp](http://www.naeyc.org/about/positions/ELL_Supplement.asp)

Nesher, P., & Katriel, T. (1977). A semantic analysis of addition and subtraction word problems in arithmetic. *Educational Studies in Mathematics*, 8, 251-269.

Oviedo, G. C. B. (2005). Comprehending algebra word problems in the first and second languages. In J. Cohen, K. T. McAlister,

K. Rolstad & J. Macswan (Eds.), *Proceedings of the 4th International Symposium on Bilingualism* (pp. 267-295). Somerville, MA: Cascadilla Press.

Rapp, J., & Allalouf, A. (2003). Evaluating cross-lingual equating. *International Journal of Testing*, 3(2), 101-117.

Reed, S.K. (1999). *Learning rules in word problems: Research and Curriculum Reform*. Lawrence Erlbaum Associates. Retrieved from [http://books.google.com.my/books?id=0FcfdR4PBHOC7dq=nctm=\(Word+Problem\)](http://books.google.com.my/books?id=0FcfdR4PBHOC7dq=nctm=(Word+Problem)).

Robertson, S., & Summerlin, J. (2005). *Mathematics and English language learners: A review of the literature*. Retrieved from <http://www.education.txsatte.edu/epic/mellwebdocs/SRSUlitreview.htm>.

Schulte, A. G., Elliott, S. N., & Kratochwill, T. R. (2001). Experimental analysis of the effects of testing accommodations on students' standardized achievement test scores. *School Psychology Review*, 30(4), 527-547.

Sireci, S. G. (1997). Problems and issues in linking tests across languages. *Journal of Educational Measurement*, 16(1), 12-19.

Stansfield, C. W. (1997). *Experience and issues related to the format of bilingual tests: Dual language test booklets versus two different test booklets*. Retrieved from <http://www.eric.ed.gov/PDFS/ED419002.pdf>

Tinkelman, S. N. (1971). Planning the objective test. In R. L. Thorndike (Ed.), *Educational measurement* (pp. 46-80). Washington, D.C.: American Council on Education.

Valentin, J.D., & Lim, C.S. (December, 2004). *Roles of semantic structure of arithmetic word problems on pupil's ability to identify the correct operation*. Retrieved from <http://www.cimt.plymouth.ac.uk/journal/valentin.pdf>

Virginia Department of Education. (2004). *Mathematics: Strategies for teaching limited English proficient students*. Retrieved from <http://www.pen.k12.va.us/VDOE/Instruction/ESL/LEP/mathResource.pdf>

Wang, M. C., Reynolds, M. C., & Walberg, H. J. (1991). *Emerging programmes: Handbook of special education research and practice*, 4. New York: Pergamon Press.

Wilde, J. (2007). *Resources about accommodations for English language learners*. National Clearinghouse for English language acquisition and language instruction educational programs. Retrieved from <http://www.nclena.gwu.edu/resabout/accomodations/intro/index.html>